

Scaling Laws for Maximum Coloring of Random Geometric Graphs

Sem Borst*

Milan Bradonjić†

November 17, 2016

Abstract

We examine maximum vertex coloring of random geometric graphs, in an arbitrary but fixed dimension, with a constant number of colors. Since this problem is neither scale-invariant nor smooth, the usual methodology to obtain limit laws cannot be applied. We therefore leverage different concepts based on subadditivity to establish convergence laws for the maximum number of vertices that can be colored. For the constants that appear in these results, we provide the exact value in dimension one, and upper and lower bounds in higher dimensions.

1 Introduction

We examine maximum coloring of random geometric graphs (RGGs), in an arbitrary but fixed dimension d , with a constant number of colors. The vertices of an RGG (whose spatial distribution will be defined below) are embedded in an Euclidean space that is equipped with the ℓ_2 distance or some ℓ_p distance in general, and two vertices are connected if and only if they are within a given Euclidean distance r . More specifically, we address the questions: *What is the maximum number of vertices in a sparse RGG that can be properly colored with a constant number of colors? In particular, what is the asymptotic behavior of that value, as the total number of vertices in the graph tends to infinity?*

It is important to emphasize the distinction between our problem and that of determining the chromatic number, which is the minimum required number of colors to properly color all the vertices of a graph such that no two adjacent vertices are assigned the same color. Determining whether an RGG (or a unit-disk graph) is k -colorable, i.e., whether its chromatic number is at most k , is NP-hard even for $k = 3$, see [3]. Our problem is different from determining the chromatic number, since we are interested in the maximum number of vertices that can be properly colored with given $k \in \mathbb{N}$ colors, as well as from k -colorability, which is a binary decision problem. The chromatic number of RGGs has been studied in detail (for different values of the expected degree), see Theorem 1.1 in [8]. The chromatic number in the *thermodynamic regime*, when the expected degree is constant, is ‘almost’ logarithmic in the number of vertices n , i.e., $(1 + o(1)) \log n / \log \log n$, which additionally inspires our problem where only a constant number of colors is available.

The above-mentioned questions are not only of fundamental interest, but also motivated by applications in wireless networks, where the various users need to be assigned channels (transmission frequencies) in order to be able to communicate, subject to certain interference constraints. For example, in order to avoid excessive interference, the same channel cannot be assigned to two users within a certain reuse distance r . The total number of required channels to cover all users then corresponds to the chromatic number

*Mathematics of Systems, Nokia Bell Labs, 600 Mountain Avenue, Murray Hill NJ 07974, USA, and Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB, The Netherlands. E-mail: sem@research.bell-labs.com.

†Mathematics of Systems, Nokia Bell Labs, 600 Mountain Avenue, Murray Hill NJ 07974, USA. E-mail: milan@research.bell-labs.com.

of the associated interference graph where two users are neighbors when they are located within distance r . When the user locations are governed by a spatial Poisson process, the interference graph is an RGG, and the chromatic number will grow without bound as the total number of users grows large. As a result, the required number of channels to cover all users will grow without bound, implying that the capacity per channel, and hence the so-called max-min throughput of the network, will vanish in the limit, which is obviously undesirable. The question thus arises how many users can be covered when the number of available channels is finite. It will then not be feasible to cover all users as the total number of users grows large, but the users that do get covered are ensured to receive a strictly positive throughput. The results that we prove in the present paper imply that any target for the fraction of users to be covered, arbitrarily close to one, can be achieved in the limit with a sufficiently large but constant number of channels.

Besides wireless networks, RGGs have also found applications in various further areas, e.g. cluster analysis, statistical physics, modeling data in high-dimensional spaces, and hypothesis testing, to mention just a few [10]. For problems on many of these ‘real’ networks, the sparse regime with constant expected vertex degree is particularly relevant, see [7].

We now formally state the main problem and results. For any subset of points $V \subseteq \mathbb{R}^d$ and $r \in \mathbb{R}_+$, let $G_r(V)$ be the graph with vertex set V and edge set $E = \{\{u, v\} \in V^2 : \|u - v\| \leq r\}$, i.e., connecting all pairs of points that are within a given Euclidean distance r .

The main object of interest is the cardinality of a set obtained by a maximum proper coloring with k colors of a given graph $G_r(V)$. Note that such a set obtained by a maximum proper coloring on finite V may not be unique, but its cardinality is unique and defined as follows. For any $k \in \mathbb{N}$, let $N_{k,r}(V)$ be the maximum number of vertices that can be properly colored in $G_r(V)$ with k colors. For any $\lambda > 0$, let \mathcal{X}_λ be a Poisson point process of intensity λ in \mathbb{R}^d . For compactness, denote

$$F_{k,\lambda}(t) = N_{k,1}([0, t]^d \cap \mathcal{X}_\lambda)$$

for any $t \geq 0$.

Also, let \mathcal{I}_n be a collection of n points uniformly and independently distributed in the unit cube $[0, 1]^d$. For compactness, denote

$$H_{k,r}(n) = N_{k,r}(\mathcal{I}_n)$$

for any $n \in \mathbb{N}$ and $r > 0$.

The main problem: We are interested in the asymptotic behavior of the expectation and moreover the distribution of $F_{k,\lambda}(t)$ as $t \rightarrow \infty$, as well as $H_{k,\nu}(n)$ as $n \rightarrow \infty$.

The main results: We show that for any $d, k \in \mathbb{N}$ and $\lambda > 0$, the functional $F_{k,\lambda}(t)$ converges in probability

$$\frac{F_{k,\lambda}(t)}{\lambda t^d} \xrightarrow{p} a_{k,\lambda},$$

for some $a_{k,\lambda} \in (0, 1]$, and in distribution, for any $\nu > 0$,

$$\frac{H_{k,\sqrt[d]{\nu/n}}(n)}{n} \xrightarrow{d} a_{k,\nu}.$$

One of our main methods involves the notion of *subadditivity*. Concretely, we divide the cube $[0, t]^d$ into cubes of volume s^d , for some $s < t$ which we specify later, and apply the subadditivity argument in order to relate $F_{k,\lambda}(t)$ and $F_{k,\lambda}(s)$. We show that the lower and upper limits as $t \rightarrow \infty$ of $F_{k,\lambda}(t)$ exist and are the same, and moreover we establish the weak law of large numbers for $F_{k,\lambda}(t)$ and the strong law of large numbers for $H_{k,\nu}(n)$.

In Lemma 4, we prove that the variance $\text{Var} \{F_{k,\lambda}(t)\} = \Omega(t^d)$, i.e., the limiting variance normalized by t^d is bounded away from 0, and in Lemma 6 we present an upper bound on $\text{Var} \{F_{k,\lambda}(t)\} = O(t^d)$, which together imply $\text{Var} \{F_{k,\lambda}(t)\} = \Theta(t^d)$, see Lemma 7.

There are two branches of methods prevalent in discrete stochastic geometry, subadditive and stabilization methods, usually used to obtain the limiting behavior of some Euclidean functionals: laws of large numbers, central limit theorems, etc. For an excellent survey, the reader is referred to Yukich [14].

At first glance, the results in this paper can be seen as a subproblem and amenable to analysis by using techniques from [2], and even the “more general subadditive methods” developed by Steele in Chapter 3 of [13]. In order to apply these techniques from [13], a function L that maps a finite subset of points from \mathbb{R}^d to \mathbb{R}_+ must satisfy the following four hypotheses: (i) normalization $L(\emptyset) = 0$; (ii) homogeneity $L(\alpha x_1, \alpha x_2, \dots, \alpha x_n) = \alpha L(x_1, x_2, \dots, x_n)$ for every $\alpha > 0$; (iii) translation invariance $\forall y \in \mathbb{R}^d$ $L(x_1 + y, x_2 + y, \dots, x_n + y) = L(x_1, x_2, \dots, x_n)$; (iv) geometric subadditivity, where for all $m, n \geq 1$ and $x_1, x_2, \dots, x_n \in [0, 1]^d$ we have

$$L(x_1, x_2, \dots, x_n) \leq \sum_{i=1}^{m^d} L(\{x_1, x_2, \dots, x_n\} \cap Q_i) + O(m^{d-1}), \quad (1)$$

where the unit cube $[0, 1]^d$ is partitioned into m^d cubes Q_i with side $1/m$.

Additionally, (v) L is monotone, if for all n and x_i , $L(x_1, \dots, x_n) \leq L(x_1, \dots, x_n, x_{n+1})$. For example, Steele proves a so-called “basic theorem”, see Theorem 3.1.1 in [13], for general subadditive Euclidean functionals that are monotone and satisfy the four conditions (i)–(iv). This theorem states that if x_1, x_2, \dots, x_n are independent random variables uniformly distributed on $[0, 1]^d$ then with probability one

$$\lim_{n \rightarrow \infty} n^{-(d-1)/d} L(x_1, x_2, \dots, x_n) = \beta_L(d), \quad (2)$$

where $\beta_L(d)$ is a positive constant, which depends both on the functional L and the dimension d .

In order to show applications of Theorem 3.1.1 from [13], we first mention, one of the classical problems in combinatorial optimization, the famous traveling salesman problem (TSP), where the goal is to find the minimum-length Hamiltonian tour among a list of cities, given the distances between each pair of cities. For more details on TSP, the reader is referred to [1].

One subcase of TSP is Euclidean TSP, where the objective is to determine the minimum-length Hamiltonian tour of n points x_1, x_2, \dots, x_n distributed in \mathbb{R}^d , and analyze its limiting behavior as n grows. Euclidean TSP was originally studied and the limiting behavior (2) was shown by Beardwood, Halton and Hammersley in [2].

Euclidean TSP is one representative example to apply Theorem 3.1.1 from [13]. It is clear that the length of the minimum tour satisfies (i)–(v). In other words, the minimum-length Hamiltonian tour on n points: (i) is equal to 0 if the number of points $n = 0$ (normalization); (ii) increases α times, for every $\alpha > 0$, if the positions of the points are rescaled α times (homogeneity); (iii) is translation invariant (does not change under the translation of the points); (iv) is subadditive, see [13, 5]; and (v) is monotone by applying the triangle inequality on three distances among x_{n+1} and its two adjacent points in the optimal tour on x_1, x_2, \dots, x_{n+1} . Besides Euclidean TSP, subadditive Euclidean functionals, such as, the length of the least Euclidean matching, Steiner tree, and rectilinear Steiner tree, satisfy the conditions of Theorem 3.1.1 from [13], and their asymptotic growth rates have been studied in [12].

However, the functionals of interest in this work $F_{k,\lambda}(t)$ and $H_{k,\nu}(n)$ satisfy normalization, translation invariance, and subadditivity, but *not* the homogeneity property, i.e., $N_{k,r}(\alpha V) \neq \alpha N_{k,r}(V)$, where $\alpha > 0$ and $V \subset \mathbb{R}^d$.

Moreover, we later show that $N_{k,r}(V)$ is *not continuous*, also called *not smooth*. In more detail, for any V and V' at distance at least r , $|N_{k,r}(V \cup V') - N_{k,r}(V)| = |N_{k,r}(V')| = \Theta(V')$, which is not $O(|V'|^{(d-p)/d})$ for some $p > 0$, hence one cannot apply the methods from [11]. For example, TSP and MST (Minimum Spanning Tree) are smooth functionals. In [14], Yukich presents laws of large numbers for smooth, superadditive Euclidean

functionals, see Theorem 8.1, as well as the general ‘umbrella theorem’ for subadditive, smooth Euclidean functionals, see Theorem 8.3.

Additionally, there is a large body of work, which uses the *stabilization* method, see Penrose [10], Yukich [14]. It is not hard to show that stabilization holds in the sub-critical regime for $\lambda < \lambda_c$, when a.s.¹ no infinite cluster exists, but it is not clear at all whether stabilization holds in the super-critical regime $\lambda > \lambda_c$ that we are interested in. Moreover, the stabilization is a very ‘effective’ method, see Penrose [10], which needs assumptions on a considered functional that seem rather strong and that we do not impose in our proofs.

2 Main Properties and Preliminary Results

As introduced before, for any subset of points $V \subseteq \mathbb{R}^d$ and $r \in \mathbb{R}_+$, let $G_r(V)$ be the graph with vertex set V and edge set $E = \{\{u, v\} \in V^2 : \|u - v\| \leq r\}$, i.e., connecting all pairs of points that are within a given distance r . For any $k \in \mathbb{N}$, let $N_{k,r}(V)$ be the maximum number of vertices that can be properly colored in $G_r(V)$ with k colors.

We state a few basic but useful properties.

- (0) $N_{k,r}(\alpha V) = N_{k,r/\alpha}(V)$, for all $\alpha > 0$, since the edge sets of $G_r(\alpha V)$ and $G_{r/\alpha}(V)$ coincide.
- (i) Inhomogeneity (*no scale-invariance*): There exists $\alpha > 0$ such that $N_{k,r}(\alpha V) \neq \alpha N_{k,r}(V)$, and in general the inequality holds for many values $\alpha > 0$.
- (ii) Monotonicity: If $U \subseteq V$, then $N_{k,r}(U) \leq N_{k,r}(V)$. Also, $N_{k,r}(V)$ is decreasing in r .
- (iii) Subadditivity: $N_{k,r}(V_1 \cup V_2) \leq N_{k,r}(V_1) + N_{k,r}(V_2)$ for any V_1, V_2 .
- (iv) Near-superadditivity: If $V_1, V_2 \subseteq V$ are at a distance more than r , i.e., $\|v_1 - v_2\| > r$ for all $v_1 \in V_1, v_2 \in V_2$, then $N_{k,r}(V_1 \cup V_2) \geq N_{k,r}(V_1) + N_{k,r}(V_2)$.

By virtue of the property (0) and the properties of the spatial Poisson process,

$$N_{k,r}([0, t]^d \cap \mathcal{X}_\lambda) \stackrel{st}{=} N_{k,1}([0, t]^d \cap \mathcal{X}_\lambda / r) \stackrel{st}{=} N_{k,1}([0, t/r]^d \cap \mathcal{X}_{\lambda r^d}) = F_{k,\lambda r^d}(t/r),$$

and hence we may focus the attention on the random variable $F_{k,\lambda}(t)$ without loss of generality.

Also, it follows from the above-mentioned monotonicity property and a simple coupling argument that $H_{k,\nu}(n)$ is stochastically increasing in n .

For any $\mu \in \mathbb{R}_+$, let $\text{Po}(\mu)$ be a Poisson random variable with parameter μ and $\pi(\mu, m) = \mathbb{P}\{\text{Po}(\mu) = m\}$. The property (0) and the properties of the spatial Poisson process furnish a useful relationship between the variables $F_{k,\lambda}(t)$ and $H_{k,\nu}(n)$:

$$F_{k,\lambda}(t) \stackrel{st}{=} N_{k,1/t}([0, t]^d \cap \mathcal{X}_\lambda) / t \stackrel{st}{=} N_{k,1/t}([0, 1]^d \cap \mathcal{X}_{\lambda t^d}) \stackrel{st}{=} H_{k,1/t}(\text{Po}(\lambda t^d)),$$

and in particular,

$$\mathbb{E}\{F_{k,\lambda}(t)\} = \mathbb{E}\{H_{k,1/t}(\text{Po}(\lambda t^d))\} = \sum_{l=0}^{\infty} \mathbb{E}\{H_{k,1/t}(l)\} \pi(\lambda t^d, l). \quad (3)$$

For any $s, t \in \mathbb{R}_+$, define

$$M^u(s, t) = \left\lceil \frac{t}{s} \right\rceil^d,$$

and

$$M^l(s, t) = \left\lfloor \frac{t}{s+1} \right\rfloor^d.$$

Invoking the stationarity (translation invariance) of the spatial Poisson process, the subadditivity property implies

$$F_{k,\lambda}(t) \leq_{st} \sum_{i=1}^{M^u(s,t)} N_i(s), \quad (4)$$

¹a.s. stands for “asymptotically almost surely”.

while the near-superadditivity property implies

$$F_{k,\lambda}(t) \geq_{st} \sum_{i=1}^{M^l(s,t)} N_i(s), \quad (5)$$

where $N_1(s), N_2(s), \dots$ are i.i.d. copies of the random variable $F_{k,\lambda}(s)$.

For compactness, denote $\overline{F}_{k,\lambda}(t) = \mathbb{E} \left\{ \tilde{F}_{k,\lambda}(t) \right\}$, with

$$\tilde{F}_{k,\lambda}(t) = \frac{F_{k,\lambda}(t)}{\lambda t^d},$$

as the ‘coloring ratio’, i.e., the ratio of the expected maximum number of vertices that can be colored and the expected total number of vertices.

The stochastic upper and lower bounds (4) and (5) yield

$$\tilde{F}_{k,\lambda}(t) \leq_{st} \left(\frac{s}{t} \right)^d \sum_{i=1}^{M^u(s,t)} \tilde{N}_i(s), \quad (6)$$

and

$$\tilde{F}_{k,\lambda}(t) \geq_{st} \left(\frac{s}{t} \right)^d \sum_{i=1}^{M^l(s,t)} \tilde{N}_i(s), \quad (7)$$

where $\tilde{N}_1(s), \tilde{N}_2(s), \dots$ are i.i.d. copies of the random variable $\tilde{F}_{k,\lambda}(s)$.

Proposition 1 *For any $k \geq 1$, $\lambda > 0$, the limit $\lim_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t)$ exists, and equals $a_{k,\lambda} := \inf_{t>0} \overline{F}_{k,\lambda}(t)$.*

Proof Taking expectations in the stochastic upper bound (6), we obtain

$$\overline{F}_{k,\lambda}(t) \leq \left(\frac{s}{t} \right)^d M^u(s,t) \overline{F}_{k,\lambda}(s).$$

Note that for any fixed s ,

$$\lim_{t \rightarrow \infty} \left(\frac{s}{t} \right)^d M^u(s,t) = 1,$$

i.e., for any $\delta > 0$, there exists $t(s, \delta)$ such that

$$\left(\frac{s}{t} \right)^d M^u(s,t) \leq 1 + \delta$$

for all $t \geq t(s, \delta)$.

This yields

$$\limsup_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t) \leq (1 + \delta) \overline{F}_{k,\lambda}(s)$$

for any fixed s , hence

$$\limsup_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t) \leq (1 + \delta) \inf_{s>0} \overline{F}_{k,\lambda}(s),$$

and therefore

$$\limsup_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t) \leq \inf_{s>0} \overline{F}_{k,\lambda}(s),$$

since $\delta > 0$ is arbitrary.

Because obviously

$$\liminf_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t) \geq \inf_{s>0} \overline{F}_{k,\lambda}(s),$$

we deduce

$$\lim_{t \rightarrow \infty} \overline{F}_{k,\lambda}(t) = \inf_{s>0} \overline{F}_{k,\lambda}(s).$$

□

The next proposition provides two useful properties for the constant $a_{k,\lambda}$ introduced in Proposition 1.

Proposition 2 *For any $\epsilon > 0$,*

$$a_{k,\lambda(1+\epsilon)} \leq a_{k,\lambda} + \epsilon,$$

and thus

$$a_{k,\lambda(1-\epsilon)} \geq a_{k,\lambda} - \frac{\epsilon}{1-\epsilon}.$$

Proof By virtue of the subadditivity property and the properties of the spatial Poisson process, for any $\epsilon > 0$, $t > 0$,

$$\begin{aligned} F_{k,\lambda(1+\epsilon)}(t) &= N_{k,1}([0, t]^d \cap \mathcal{X}_{\lambda(1+\epsilon)}) \\ &\stackrel{st}{=} N_{k,1}([0, t]^d \cap \mathcal{X}_\lambda \cup ([0, t]^d \cap \mathcal{X}'_{\epsilon\lambda})) \\ &\leq_{st} N_{k,1}([0, t]^d \cap \mathcal{X}_\lambda) + N_{k,1}([0, t]^d \cap \mathcal{X}'_{\epsilon\lambda}) \\ &\leq N_{k,1}([0, t]^d \cap \mathcal{X}_\lambda) + |[0, t]^d \cap \mathcal{X}'_{\epsilon\lambda}|. \end{aligned}$$

Taking expectations and dividing by λt^d yields

$$\bar{F}_{k,\lambda(1+\epsilon)}(t) \leq \bar{F}_{k,\lambda}(t) + \epsilon.$$

Taking the limit over $t > 0$, we obtain the statement of the proposition.

□

For any $k \geq 1$, $\nu > 0$, define $\tilde{H}_{k,\nu}(n) = H_{k,\nu}(n)/n$.

Proposition 3 *For any $k \geq 1$, $\nu > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left\{ \tilde{H}_{k, \sqrt[d]{\nu/n}}(n) \right\} = a_{k,\nu}.$$

Proof The proof relies on the relationship (3) and the monotonicity property of $H_{k,\nu}(n)$, in conjunction with asymptotic lower and upper bounds.

We first consider the upper bound.

Using (3) and observing that $\mathbb{E} \{H_{k,1/t}(l)\}$ is increasing in l , we obtain

$$\mathbb{E} \{F_{k,\lambda}(t)\} \geq \mathbb{E} \{H_{k,1/t}(n)\} \mathbb{P} \{\text{Po}(\lambda t^d) \geq n\},$$

for any $n \in \mathbb{N}$.

Taking $t = \sqrt[d]{n/\nu}$ and $\lambda = \nu(1+\epsilon)$, we deduce

$$\mathbb{E} \left\{ F_{k,\nu(1+\epsilon)}(\sqrt[d]{n/\nu}) \right\} \geq \mathbb{E} \left\{ H_{k,\sqrt[d]{\nu/n}}(n) \right\} \mathbb{P} \{\text{Po}(n(1+\epsilon)) \geq n\}.$$

Noting that $\mathbb{P} \{\text{Po}(n(1+\epsilon)) \geq n\} \rightarrow 1$ as $n \rightarrow \infty$, Proposition 1 then implies

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E} \left\{ H_{k,\sqrt[d]{\nu/n}}(n) \right\}}{n} \leq (1+\epsilon) \limsup_{n \rightarrow \infty} \frac{\mathbb{E} \left\{ F_{k,\nu(1+\epsilon)}(\sqrt[d]{n/\nu}) \right\}}{\nu(1+\epsilon)n/\nu} = (1+\epsilon)a_{k,\nu(1+\epsilon)}.$$

Since $\epsilon > 0$ is arbitrary, it then follows from Proposition 2 that

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E} \left\{ H_{k,\sqrt[d]{\nu/n}}(n) \right\}}{n} \leq a_{k,\nu}.$$

We now turn to the lower bound.

Invoking (3) and noting that $\mathbb{E}\{H_{k,1/t}(l)\}$ is increasing in l , we obtain

$$\mathbb{E}\{F_{k,\lambda}(t)\} \leq \mathbb{E}\{H_{k,1/t}(n)\} + \sum_{l=n}^{\infty} l\mathbb{P}\{\text{Po}(\lambda t^d) = l\},$$

for any $n \in \mathbb{N}$.

Choosing $t = \sqrt[d]{n/\nu}$ and $\lambda = \nu(1 - \epsilon)$, we deduce

$$\mathbb{E}\{F_{k,\nu(1-\epsilon)}(\sqrt[d]{n/\nu})\} \leq \mathbb{E}\{H_{k,\sqrt[d]{\nu/n}}(n)\} + \sum_{l=n}^{\infty} l\mathbb{P}\{\text{Po}(n(1-\epsilon)) = l\}.$$

Invoking that $\sum_{l=n}^{\infty} l\mathbb{P}\{\text{Po}(n(1-\epsilon)) = l\} = o(1)$ as $n \rightarrow \infty$, it then follows from Proposition 1 that

$$\liminf_{n \rightarrow \infty} \frac{\mathbb{E}\{H_{k,\sqrt[d]{\nu/n}}(n)\}}{n} \geq (1 - \epsilon) \liminf_{n \rightarrow \infty} \frac{\mathbb{E}\{F_{k,\nu(1-\epsilon)}(\sqrt[d]{n/\nu})\}}{\nu(1 - \epsilon)n/\nu} = (1 - \epsilon)a_{k,\nu(1-\epsilon)}.$$

Since $\epsilon > 0$ is arbitrary, it then follows from Proposition 2 that

$$\liminf_{n \rightarrow \infty} \frac{\mathbb{E}\{H_{k,\sqrt[d]{\nu/n}}(n)\}}{n} \geq a_{k,\nu}.$$

□

For conciseness, denote $\mu(t) = \mathbb{E}\{F_{k,\lambda}(t)\}$ and $\sigma^2(t) = \mathbb{V}ar\{F_{k,\lambda}(t)\}$.

Lemma 4 For all λ, d, k ,

$$\inf_{t \geq 3} \frac{\sigma^2(t)}{t^d} > 0.$$

Proof Let $C^d(t) = (\lfloor \frac{t}{3} \rfloor)^d$ and let $\mathcal{V} = v_1, v_2, \dots, v_{C^d(t)}$ be a set of points in $[0, t]^d$ such that no two points are within distance 3 and no point is within distance $3/2$ from the boundary. For each point v_i consider two spheres with radii $1/2$ and $3/2$ centered on v_i , along with the shell of unit width consisting of the area covered by the larger sphere but not by the smaller sphere. Let $\mathcal{W} \subseteq \mathcal{V}$ be the set of points for which the associated shells do not contain any points of the Poisson point process \mathcal{X}_λ . Note that

$$\mathbb{P}\{\mathcal{W}\} = q^{|\mathcal{W}|}(1 - q)^{C^d(t) - |\mathcal{W}|},$$

with

$$q = e^{-\lambda(V^d(3/2) - V^d(1/2))} > 0,$$

and

$$V^d(r) = \frac{\pi^{d/2} r^d}{\Gamma(d/2 + 1)}$$

denoting the volume of a sphere of radius r in d dimensions. Given a maximum proper coloring with k colors of $[0, t]^d \cap \mathcal{X}_\lambda$, let $D_i \in \{0, 1, \dots, k\}$ be the number of colored points covered by the smaller sphere around v_i and $D_0 = F_{k,\lambda}(t) - \sum_{i \in \mathcal{W}} D_i$. We may write

$$\sigma^2(t) = \mathbb{E}\{(F_{k,\lambda}(t) - \mu(t))^2\} = \sum_{\mathcal{W} \subseteq \{1, \dots, C^d(t)\}} \mathbb{E}\{(F_{k,\lambda}(t) - \mu(t))^2 | \mathcal{W}\} \mathbb{P}\{\mathcal{W}\}.$$

Now observe that

$$D_i \stackrel{st}{=} \min\{k, \text{Po}(\lambda V^d(1/2))\},$$

for all $i \in \mathcal{W}$, and that D_0 and D_i , $i \in \mathcal{W}$, are all mutually independent conditioned on \mathcal{W} . Denoting $\mu_k = \mathbb{E} \{ \min\{k, \text{Po}(\lambda V^d(1/2))\} \}$, and $\sigma_k^2 = \text{Var} \{ \min\{k, \text{Po}(\lambda V^d(1/2))\} \}$, we deduce

$$\begin{aligned}
\mathbb{E} \left\{ (F_{k,\lambda}(t) - \mu(t))^2 \mid \mathcal{W} \right\} &= \mathbb{E} \left\{ \left(D_0 + \sum_{i \in \mathcal{W}} D_i - \mu(t) \right)^2 \mid \mathcal{W} \right\} \\
&= \mathbb{E} \left\{ \left(D_0 - (\mu(t) - |\mathcal{W}| \mu_k) + \sum_{i \in \mathcal{W}} (D_i - \mu_k) \right)^2 \mid \mathcal{W} \right\} \\
&= \mathbb{E} \left\{ (D_0 - (\mu(t) - |\mathcal{W}| \mu_k))^2 \mid \mathcal{W} \right\} + \sum_{i \in \mathcal{W}} \mathbb{E} \left\{ (D_i - \mu_k)^2 \mid \mathcal{W} \right\} \\
&\geq |\mathcal{W}| \sigma_k^2.
\end{aligned}$$

Combining the above properties yields

$$\sigma^2(t) \geq \sum_{\mathcal{W} \subseteq \{1, \dots, C^d(t)\}} |\mathcal{W}| \sigma_k^2 q^{|\mathcal{W}|} (1-q)^{C^d(t)-|\mathcal{W}|} = \sigma_k^2 \sum_{m=0}^{C^d(t)} \binom{C^d(t)}{m} m q^m (1-q)^{C^d(t)-m} = q C^d(t) \sigma_k^2.$$

Since $q > 0$, $\sigma_k^2 > 0$, regardless of t , and $C^d(t) \geq (t/6)^d$ for all $t > 3$, the statement of the lemma follows with

$$\inf_{t > 3} \frac{\sigma^2(t)}{t^d} \geq q \sigma_k^2 / 6^d > 0. \quad (8)$$

□

Lemma 5 For any $n \in \mathbb{N}$,

$$\text{Var} \left\{ H_{k, \sqrt[d]{\nu/n}}(n) \right\} \leq n/2.$$

Proof Recall that \mathcal{I}_n is a collection of n points uniformly and independently distributed in the unit cube $[0, 1]^d$. Let \mathcal{I}_n^i be a collection of n points obtained from \mathcal{I}_n by replacing (only) the i -th point x_i by $x'_i \in [0, 1]^d$. By definition, $H_{k,r}(n) = N_{k,r}(\mathcal{I}_n)$. Moreover, $N_{k,r}(\mathcal{I}_n)$ is one-Lipschitz

$$|N_{k,r}(\mathcal{I}_n) - N_{k,r}(\mathcal{I}_n^i)| \leq 1, \quad (9)$$

by the nature of a maximum proper coloring with k colors. Now the one-Lipschitz property of $N_{k,r}(\mathcal{I}_n)$ and the Efron-Stein bound [4] yield the assertion of the lemma

$$\begin{aligned}
\text{Var} \left\{ H_{k, \sqrt[d]{\nu/n}}(n) \right\} &= \text{Var} \left\{ N_{k, \sqrt[d]{\nu/n}}(\mathcal{I}_n) \right\} \\
&\leq \frac{1}{2} \mathbb{E} \left\{ \sum_{i=1}^n \left(N_{k, \sqrt[d]{\nu/n}}(\mathcal{I}_n) - N_{k, \sqrt[d]{\nu/n}}(\mathcal{I}_n^i) \right)^2 \right\} \\
&\leq n/2.
\end{aligned}$$

□

Lemma 6 For any λ, d, k and for every $\varepsilon > 0$ and sufficiently large t ,

$$\sigma^2(t) \leq (1 + \varepsilon) \lambda t^d / 2.$$

Proof Conditioning on the number of points in $\mathcal{X}_\lambda \cap [0, t]^d$, which is a Poisson random variable and concentrates around its mean λt^d , see [10], the proof follows by using the bound on the variance of $H_{k, \sqrt[d]{\nu/n}}(n)$ provided by Lemma 5.

□

Lemma 7 For all $\lambda > 0$ and $d, k \in \mathbb{N}$, asymptotically as t tends to ∞ , we have

$$\sigma^2(t) = \Theta(t^d).$$

Proof The proof follows from Lemmas 4 and 6. □

3 Convergence of functionals $\tilde{F}_{k,\lambda}(t)$ and $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$

The next two propositions establish convergence in probability, i.e., a weak law-of-large-numbers result for the two functionals $\tilde{F}_{k,\lambda}(t)$ and $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$, and a strong law-of-large-numbers result (almost-sure convergence) of $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$.

Proposition 8 For any $k \geq 1$, $\nu > 0$, the random variable $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$ almost surely converges to $a_{k,\nu}$ as $n \rightarrow \infty$.

The first part of the proof below establishes the convergence in probability, while the second part concludes with the almost-sure convergence.

Proof We first show the convergence in probability. We have that $N_{k,r}(V)$ is one-Lipschitz, see 9. Hence a ‘simple concentration bound’, see p.79 (Section 10.1) of [9], implies that for any $\varepsilon \geq 0$

$$\mathbb{P} \left\{ \left| H_{k,\sqrt[d]{\nu/n}}(n) - \mathbb{E} \left\{ H_{k,\sqrt[d]{\nu/n}}(n) \right\} \right| \geq \varepsilon \right\} \leq 2e^{-\varepsilon^2/2n}.$$

Moreover, for any $\delta > 0$, Proposition 3 implies that there exists an $n = n_\delta$ such that

$$\left| \mathbb{E} \left\{ \tilde{H}_{k,\sqrt[d]{\nu/n}}(n) \right\} - a_{k,\lambda} \right| \leq \frac{1}{2}\delta$$

for all $n \geq n_\delta$. Thus, for all $n \geq n_\delta$,

$$\begin{aligned} \mathbb{P} \left\{ \left| \tilde{H}_{k,\sqrt[d]{\nu/n}}(n) - a_{k,\nu} \right| > \delta \right\} &\leq \mathbb{P} \left\{ \left| \tilde{H}_{k,\sqrt[d]{\nu/n}}(n) - \mathbb{E} \left\{ \tilde{H}_{k,\sqrt[d]{\nu/n}}(n) \right\} \right| > \frac{1}{2}\delta \right\} \\ &= \mathbb{P} \left\{ \left| H_{k,\sqrt[d]{\nu/n}}(n) - \mathbb{E} \left\{ H_{k,\sqrt[d]{\nu/n}}(n) \right\} \right| > \frac{n}{2}\delta \right\} \\ &\leq 2e^{-n\delta^2/2}. \end{aligned} \tag{10}$$

For any $\delta > 0$, the latter term tends to 0 as n tends to ∞ , which implies the convergence in probability of $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$.

We now proceed to establish the almost-sure convergence. From (10), for any $\delta > 0$ and the corresponding finite n_δ , it follows:

$$\sum_{n=1}^{\infty} \mathbb{P} \left\{ \left| \tilde{H}_{k,\sqrt[d]{\nu/n}}(n) - a_{k,\nu} \right| > \delta \right\} \leq n_\delta + 2 \sum_{n=n_\delta}^{\infty} e^{-n\delta^2/2} < \infty, \tag{11}$$

which implies the almost-sure convergence of $\tilde{H}_{k,\sqrt[d]{\nu/n}}(n)$. □

Given Proposition 8, we prove the convergence in probability of $\tilde{F}_{k,\lambda}(t)$.

Proposition 9 For any $k \geq 1$, $\lambda > 0$, the random variable $\tilde{F}_{k,\lambda}(t)$ converges to $a_{k,\lambda}$ in probability as $t \rightarrow \infty$.

Proof The proof relies on lower and upper bounds, which asymptotically coincide.

For both bounds, we will use the fact that

$$\mathbb{V}ar \left\{ \tilde{F}_{k,\lambda}(s) \right\} \leq \frac{1}{\lambda s^d}$$

because of Lemma 6, and hence

$$\mathbb{P} \left\{ \left| \frac{1}{M} \sum_{i=1}^M \tilde{N}_i(s) - \mathbb{E} \left\{ \tilde{F}_{k,\lambda}(s) \right\} \right| \geq \varepsilon \right\} \leq \frac{1}{M \lambda s^d \varepsilon^2} \quad (12)$$

for any $M \geq 1$ and $\varepsilon > 0$ by virtue of Chebyshev's inequality.

We first consider the lower bound.

Proposition 1 implies that there exists an $s = s_\delta$ such that

$$\overline{F}_{k,\lambda}(s) \geq a_{k,\lambda} - \frac{1}{4}\delta.$$

There also exists a $t(s, \delta)$ so that

$$\left(\frac{t}{s} \right)^d \frac{1}{M^l(s, t)} \leq 1 + \frac{1}{4}\delta$$

for all $t \geq t(s, \delta)$.

Noting that $\overline{F}_{k,\lambda}(s) \leq 1$ and $M^l(s, t) \leq \left(\frac{t}{s} \right)^d$, we then have

$$\begin{aligned} \left(\frac{t}{s} \right)^d \frac{1}{M^l(s, t)} (a_{k,\lambda} - \delta) &\leq \left(\frac{t}{s} \right)^d \frac{1}{M^l(s, t)} \left(\overline{F}_{k,\lambda}(s) - \frac{3}{4}\delta \right) \\ &\leq \left(1 + \frac{1}{4}\delta \right) \left(\overline{F}_{k,\lambda}(s) - \frac{3}{4}\delta \right) \\ &\leq \overline{F}_{k,\lambda}(s) - \frac{1}{2}\delta \\ &= \mathbb{E} \left\{ \tilde{F}_{k,\lambda}(s) \right\} - \frac{1}{2}\delta \end{aligned}$$

for all $t \geq t(s, \delta)$.

Using the stochastic lower bound (5), the above inequality and (12), we derive for all $t \geq t(s, \delta)$,

$$\begin{aligned} \mathbb{P} \left\{ \tilde{F}_{k,\lambda}(t) \leq a_{k,\lambda} - \delta \right\} &\leq \mathbb{P} \left\{ \left(\frac{s}{t} \right)^d \sum_{i=1}^{M_t(s,t)} \tilde{N}_i(s) \leq a_{k,\lambda} - \delta \right\} \\ &= \mathbb{P} \left\{ \frac{1}{M^l(s, t)} \sum_{i=1}^{M^l(s,t)} \tilde{N}_i(s) \leq \left(\frac{t}{s} \right)^d \frac{1}{M^l(s, t)} (a_{k,\lambda} - \delta) \right\} \\ &\leq \mathbb{P} \left\{ \frac{1}{M^l(s, t)} \sum_{i=1}^{M^l(s,t)} \tilde{N}_i(s) \leq \mathbb{E} \left\{ \tilde{F}_{k,\lambda}(s) \right\} - \frac{1}{2}\delta \right\} \\ &\leq \frac{4}{\lambda s^d M^l(s, t) \delta^2}. \end{aligned}$$

The latter term tends to 0 as t grows large for any $\delta > 0$.

We now turn to the upper bound.

Proposition 1 implies that there exists an $s = s_\delta$ such that

$$\overline{F}_{k,\lambda}(s) \leq a_{k,\lambda} + \frac{1}{4}\delta.$$

There also exists a $t(s, \delta)$ so that

$$\left(\frac{t}{s}\right)^d \frac{1}{M^u(s, t)} \geq 1 - \frac{1}{8}\delta$$

for all $t \geq t(s, \delta)$.

Noting that $\bar{F}_{k,\lambda}(t) \leq 1$ and assuming $\delta \leq 1$, we then have

$$\begin{aligned} \left(\frac{t}{s}\right)^d \frac{1}{M^u(s, t)} (a_{k,\lambda} + \delta) &\geq \left(\frac{t}{s}\right)^d \frac{1}{M^u(s, t)} (\bar{F}_{k,\lambda}(s) + \frac{3}{4}\delta) \geq \left(1 - \frac{1}{8}\delta\right) (\bar{F}_{k,\lambda}(s) + \frac{3}{4}\delta) \geq \\ &\bar{F}_{k,\lambda}(s) + \frac{5}{8}\delta - \frac{3}{32}\delta^2 \geq \bar{F}_{k,\lambda}(s) + \frac{1}{2}\delta = \mathbb{E}\{\tilde{F}_{k,\lambda}(s)\} + \frac{1}{2}\delta \end{aligned}$$

for all $t \geq t(s, \delta)$.

Using the stochastic upper bound (6), the above inequality and (12), we derive for all $t \geq t(s, \delta)$,

$$\begin{aligned} \mathbb{P}\{\tilde{F}_{k,\lambda}(t) \geq a_{k,\lambda} + \delta\} &\leq \mathbb{P}\left\{\left(\frac{s}{t}\right)^d \sum_{i=1}^{M_u(s,t)} \tilde{N}_i(s) \geq a_{k,\lambda} + \delta\right\} \\ &= \mathbb{P}\left\{\frac{1}{M^u(s, t)} \sum_{i=1}^{M^u(s,t)} \tilde{N}_i(s) \geq \left(\frac{t}{s}\right)^d \frac{1}{M^u(s, t)} (a_{k,\lambda} + \delta)\right\} \\ &\leq \mathbb{P}\left\{\frac{1}{M^u(s, t)} \sum_{i=1}^{M^u(s,t)} \tilde{N}_i(s) \geq \mathbb{E}\{\tilde{F}_{k,\lambda}(s)\} + \frac{1}{2}\delta\right\} \\ &\leq \frac{4}{\lambda s^d M^u(s, t) \delta^2}. \end{aligned}$$

The latter term tends to 0 as t grows large for any $\delta > 0$.

□

4 Evaluation of $a_{k,\lambda}$

In Propositions 8 and 9 we showed that the fraction of nodes that can be properly colored converges in probability to a constant $a_{k,\lambda}$ as the size of the area and the total number of nodes grow large. It appears difficult to obtain an explicit expression for $a_{k,\lambda}$ in general. Below we provide the exact value for the one-dimensional case $d = 1$ and present lower and upper bounds for $d > 1$.

4.1 One-dimensional case (line)

In case $d = 1$, a maximum proper coloring of points in the interval $[0, t]$ can be obtained in a greedy manner by sequential inspection of these points from left to right. Specifically, a point at position $x \in [0, t]$ is selected if fewer than k points in $[x-1, x]$ have been included, and then assigned any color that is not already used for any points in $[x-1, x]$, and skipped otherwise. If we now interpret the points in $[0, t]$ as arrival times of customers, then it can be verified that the maximum proper coloring obtained in the above fashion corresponds exactly to the arrival times of those customers that are admitted in a so-called Erlang loss system with service times 1 and capacity k , starting from an empty state at time 0. In particular, the value of $F_{k,\lambda}(t)$ equals the number of admitted customers $A(k, \lambda, t)$ in the latter Erlang loss system, where the arrivals are governed by a Poisson process. It is well known [6] that $A(k, \lambda, t)/(\lambda t)$ converges in probability to $1 - \text{Erl}(\lambda, k)$ as $t \rightarrow \infty$, where

$$\text{Erl}(\lambda, k) = \frac{\frac{\lambda^k}{k!}}{\sum_{l=0}^k \frac{\lambda^l}{l!}} = \frac{\mathbb{P}\{\text{Po}(\lambda) = k\}}{\mathbb{P}\{\text{Po}(\lambda) \leq k\}}$$

denotes the so-called Erlang loss probability. Thus

$$a_{k,\lambda} = 1 - \frac{\frac{\lambda^k}{k!}}{\sum_{l=0}^k \frac{\lambda^l}{l!}} = \frac{\sum_{m=0}^{k-1} \frac{\lambda^m}{m!}}{\sum_{l=0}^k \frac{\lambda^l}{l!}}.$$

4.2 Higher dimensions

For $d > 1$ it seems difficult to derive an explicit expression for $a_{k,\lambda}$. In order to establish bounds, let g_{\max} be the maximum density that any set of points $V \subseteq \mathbb{R}_+^d$ can have such that no two points are within distance 1. Thus there is then a sphere of unit radius around every point in V that does not contain any other point in V . Hence a trivial upper bound is $g_{\max} \leq 1/V_d(1)$, where $V_d(1)$ is the volume of a sphere of unit radius in d dimensions. Now observe that any two points that are assigned the same color cannot be within distance 1, i.e., the density of the points that are assigned the same color can be at most g_{\max} . This implies that $a_{k,\lambda} \leq \max\{1, kg_{\max}/\lambda\}$. While admittedly crude, we expect this upper bound to be asymptotically tight for large values of λ , in the sense that $\lambda a_{k,\lambda} \rightarrow kg_{\max}$ as $\lambda \rightarrow \infty$.

For a given s , let $m(s)$ be the minimum number of sets $B_1, \dots, B_{m(s)}$ needed such that (i) any point in \mathbb{R}^d is within distance s from one of the points in $B_1 \cup \dots \cup B_{m(s)}$ and (ii) no two points within each of the sets B_m are within distance $1 + 2s$, $m = 1, \dots, m(s)$. The points in $B_1 \cup \dots \cup B_{m(s)}$ may be collectively interpreted as ‘cell anchor points’ or ‘base stations’. Now suppose we partition the k colors in $m(s)$ disjoint groups $C_1, \dots, C_{m(s)}$ of size at least $\lfloor k/m(s) \rfloor$, and then allocate all the colors in the group C_m to all the points in the set B_m , $m = 1, \dots, m(s)$. In order to color a point, we will assign it one of the colors allocated to the nearest anchor point. Since any point in \mathbb{R}^d is within distance s from one of the anchor points, the points that need to be supported by a given anchor point are all within radius s , and hence the total number is bounded from above by a Poisson random variable with parameter $\lambda V_d(s)$. This implies that $a_{k,\lambda} \geq \mathbb{E} \{ \min\{\text{Po}(\lambda V_d(s)), \lfloor k/m(s) \rfloor\} \} / \lambda V_d(s)$. While rather rough, this upper bound demonstrates that $a_{k,\lambda} \rightarrow 1$ as $k \rightarrow \infty$. Thus, any target coloring ratio, however close to one, can be asymptotically achieved with a sufficiently large but constant number of colors.

5 Conclusion

We have examined maximum vertex coloring of random geometric graphs, in an arbitrary but fixed dimension. This is a problem in discrete stochastic geometry which is neither scale-invariant nor smooth, and hence the traditional methodological framework to obtain limit laws cannot be applied. We have therefore leveraged different concepts based on subadditivity to establish convergence laws for the maximum number of vertices that can be colored with a constant number of colors. For the constants involved in these results, we derived the exact value in dimension one, and upper and lower bounds in higher dimensions.

The approach that we have developed for this specific non-linear Euclidean problem, offers great potential to be extended to a broader class of problems, which will be the subject of further research. Some specific questions that we aim to address in future work are: (1) What can one say when the ‘coverage area’ is not a disk, but some arbitrarily shaped area? (2) Can one prove generalizations when the disk radius r is not a constant, but a generally distributed random variable?

References

- [1] David L. Applegate, Robert E. Bixby, Vasek Chvatal, and William J. Cook. *The Traveling Salesman Problem: A Computational Study (Princeton Series in Applied Mathematics)*. Princeton University Press, Princeton, NJ, USA, 2007.

- [2] J. Beardwood, J. H. Halton, and J. M. Hammersley. The shortest path through many points. *Proc. Cambridge Philos. Soc.*, 55:299–327, 1959.
- [3] Brent N. Clark, Charles J. Colbourn, and David S. Johnson. Unit disk graphs. *Discrete Mathematics*, 86(1-3):165–177, December 1990.
- [4] B. Efron and C. Stein. The jackknife estimate of variance. *Ann. Statist.*, 9(3):586–596, May 1981.
- [5] Michel X. Goemans and Dimitris J. Bertsimas. Probabilistic analysis of the Held and Karp lower bound for the Euclidean traveling salesman problem. *Mathematics of Operations Research*, 16(1):72–89, 1991.
- [6] Frank P. Kelly. *Reversibility and stochastic networks*. Wiley series in probability and mathematical statistics. Wiley, New York, Chichester, 1979.
- [7] Jure Leskovec, Kevin J. Lang, Anirban Dasgupta, and Michael W. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009.
- [8] Colin McDiarmid and Tobias Müller. On the chromatic number of random geometric graphs. *Combinatorica*, 31(4):423–488, 2011.
- [9] Michael Molloy and Bruce Reed. *Graph Colouring and the Probabilistic Method*. Springer-Verlag, 2002.
- [10] Mathew D. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.
- [11] WanSoo T. Rhee. A matching problem and subadditive Euclidean functionals. *Ann. Appl. Probab.*, 3(3):794–801, 1993.
- [12] J. Michael Steele. Subadditive euclidean functionals and nonlinear growth in geometric probability. *Ann. Probab.*, 9(3):365–376, 1981.
- [13] J. Michael Steele. *Probability Theory and Combinatorial Optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1997.
- [14] Joseph Yukich. Limit theorems in discrete stochastic geometry. In Evgeny Spodarev, editor, *Stochastic Geometry, Spatial Statistics and Random Fields*, volume 2068 of *Lecture Notes in Mathematics*, pages 239–275. Springer Berlin Heidelberg, 2013.